Remarks

Claims 49, 50, 52, 53, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76 and 77 are pending. Claims 51, 54 and 55 have been newly cancelled. Claims 49, 50, 52, 53, 56 and 57 have been newly amended. Claims 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76 and 77 are newly added. Support for these amendments are found throughout the specification and in the claims as originally filed. No new matter has been entered. All newly added claims are encompassed by Group I of the restriction requirement drawn to methods of identifying biomarkers for osteoarthritis and methods for diagnosis and prognosis of osteoarthritis, further restricted to a DMN gene.

Claims 50, 70 and 72 clarify that said levels of RNA encoded by said gene are in blood samples which include all of the types of leukocytes in whole blood, i.e. of blood samples which include granulocytes in addition to mononuclear cells (T-lymphocytes, B-lymphocytes and monocytes). This phrase finds clear support in the specification, including at Figure 5C which shows standardized levels of insulin gene expression in each of the fractions of leukocytes which collectively constitute unfractionated leukocytes, i.e. granulocytes, T-lymphocytes, B-lymphocytes and monocytes (labeled "G.R.", "CD 3+", "CD19" and "MONO", i.e., respectively). It is well known to the ordinarily skilled artisan that CD3 and CD19 are specific cell surface markers of T-lymphocytes and B-lymphocytes (refer, for example, to the enclosed Abstract of Casey *et al.*, 1988. simplified plastic embedding and immunohistologic technique for immunophenotypic analysis of human hematopoietic and lymphoid tissues. Am J Pathol. 131:183-9). The fact that granulocytes (G.R.), lymphocytes [T-lymphocytes (CD 3+) and B-lymphocytes (CD19+)] and monocytes (MONO) represent all of the types of leukocytes found in blood is taught at Fig. A.23 Immunobiology. Garland Publishing. 2001. Fifth Edition. Janeway, Travers, Walport, and Shlomchik, eds. (attached) which clearly teaches that leukocytes are composed of granulocytes and mononuclear cells, and that the latter are composed of lymphocytes and monocytes. Additional support for the term "leukocytes" is found at paragraphs [0004] and [0005] and [0088] of the published application (US 2004-0248170).

New independent claim 71 claims a method of classifying expression of a DMN gene in a test subject relative to a population of control subjects that includes subjects classified as having

osteoarthritis and subjects classified as healthy subjects. New claim 71 comprises a step of quantifying a level of RNA encoded by the gene in a blood sample from the test subject, and a subsequent step of comparing the level in the sample from the test subject with levels of RNA encoded by the gene in blood samples from the control subjects classified as having osteoarthritis and in blood samples from the control subjects classified as healthy subjects. The new claim concludes that a statistically significant determination that the level in the sample from the test subject is similar to the levels in the samples from the control subjects classified as having having osteoarthritis and is different relative to the levels in the samples from the control subjects classified as healthy subjects classifies the expression of the gene in test subject with that in the control subjects classified as having osteoarthritis; and that a statistically significant determination that the level in the sample from the test subject is statistically different relative to the levels in the samples from the control subjects classified as having having osteoarthritis and is statistically similar to the levels in the samples from the control subjects classified as healthy subjects classifies the expression of the gene in the test subject with that the control subjects classified as healthy subjects. Support for reciting comparison of biomarker RNA levels of a test subject with those of control subjects having a disease (i.e. osteoarthritis) and with those of healthy control subjects, and determination of a statistically significant difference or similarity therebetween can be found in the published application (US 2004-0248170), for example at paragraph [0127] ("*When comparing two or more samples for differences, results are reported as statistically significant when there is only a small probability that similar results would have been observed if the tested hypothesis (i.e., the genes are not expressed at different levels) were true*"), and at paragraph [0128] ("*When comparing two or more samples for similarities, results are reported as statistically significant when there is only a small probability that similar results would have been observed if the tested hypothesis (i.e., the genes are not expressed at different levels) were true*"), respectively. Support for reciting classification of a test subject level relative to spc control levels can be found, for example, at claim 12 as originally filed ("*d) determining whether the level of said one or more gene transcripts of step a) classify with the levels of said transcripts in step b) as compared with the levels of said transcripts in step c)*"), at paragraph [0135] (relating to "*Methods that can be used for class prediction analysis*"), paragraph [0390] ("*Blood samples were taken from patients who were diagnosed with osteoarthritis and a specific*

stage of osteoarthritis as defined herein. Gene expression profiles were then analysed and compared to profiles from patients unaffected by any disease.").

Claims Rejection - 35 U.S.C. 112 2nd

Claims 51-57 are rejected under 35 U.S.C. 112, 2nd paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter which Applicant regards as the invention.

The office action indicates that the recitation of "unfractionated samples of lysed blood" in claim 51 is unclear. Although Applicant respectfully traverses, Applicant has canceled claim 51 and dependent claims 54-55 solely for the purposes of advancing prosecution without prejudice for pursuing the unclaimed subject material in another application, rendering the rejection of claims 51 and 54-55 moot. Applicant has amended dependent claims 52, 53, 56 and 57 to be dependent only from claim 49 or 50, which do not recite the phrase "unfractionated samples of lysed blood".

Claims Rejection - 35 U.S.C. 112 1st

Claims 51-57 are rejected under 35 U.S.C. 112, 1st paragraph, as failing to comply with the written description requirement on the grounds that the instantly recited phrase "unfractionated samples of lysed blood" is new matter. Although Applicant respectfully traverses, Applicant has canceled claim 51 and dependent claims 54 and 55 solely for the purposes of advancing prosecution without prejudice for pursuing the unclaimed subject matter in another application, rendering the rejection of claims 47, 50 and 51 moot. Applicant has amended dependent claims 52, 53, 56 and 57 to be dependent only from claim 49 or 50, which do not recite the phrase "unfractionated samples of lysed blood".

Claims 49-57 are rejected under 35 U.S.C. 112, 1st paragraph, as failing to comply with the enablement requirement.

Applicant respectfully traverses. Applicant disagrees with the rejection's assertion that the skilled artisan would have required an undue amount of experimentation to make and/or use the claimed invention in view of the breadth of the claims, the amount of guidance provided by

the specification and the level of predictability in the art.

The rejected claims include the steps of determining the level of RNA encoded by a DMN gene in a blood sample obtained from a human test subject and comparing it to the level of control RNA encoded by the DMN gene in blood samples of control subjects, wherein the comparison is indicative of osteoarthritis in said human test subject.

Applicant specifically traverses the statement on page 4 of the office action that "the independent claim, as written, states that a comparison of a human test subject DMN RNA level in a blood sample to a control indicates that osteoarthritis is present in the test subject", and the statement on page 5 of the office action that the '*The claims are extremely broad because they set forth that any or all comparison between a test subject and RNA level from "control subjects" is indicative of disease*'. Applicant clarifies that the phrase "wherein said comparison of said quantified level of step (a) with said quantified level of said control subjects is indicative of osteoarthritis in said human test subject" of independent claim 49, is a narrowing limitation, limiting the claim to only those comparisons which are indicative of the test subject having osteoarthritis, and excluding those comparisons which do not indicate that the test individual has osteoarthritis.

However, in the interest of expediting prosecution, Applicant has added new claims which more clearly set forth the subject matter of the newly cancelled claims. Specific points raised in the instant enablement rejection will be addressed to the extent that they are relevant to the newly added claims.

The rejection asserts that the claims are broad with respect to "control subjects", indicating that "control subjects" could encompass patients with osteoarthritis, healthy patients, and patients with some other disease such as renal cell carcinoma or a particular stage of osteoarthritis (page 5 of the office action). The instant claims recite two clearly defined sets of controls; patients having osteoarthritis and healthy controls. For example, independent claims 49, 68 and 71, and dependent claim 65 include a limitation that the controls include healthy subjects.

The office action states that the claims do not "set forth the direction of the difference necessary to indicate osteoarthritis" (page 5 of the Office Action) and suggests that without

providing this information, the mere observation of differences is an unpredictable indicator of osteoarthritis.

The Applicant respectfully submits that the invention is taught in such terms that one skilled in the art can make and use the claimed invention, including the use of the elected biomarker DMN as an indicator of osteoarthritis as described in the claims without disclosing the direction or the level of difference that exists between patients having osteoarthritis and individuals not having osteoarthritis. The Applicant has identified the elected gene DMN as differentially expressed as between individuals diagnosed as having osteoarthritis and individuals not having osteoarthritis by demonstrating a statistical difference in the level of RNA, as described in Example 24 and Table 3O. The statistical significance of the differential expression of DMN is evidenced by its P value of 0.0000415, as listed in Table 3O. Therefore the Applicant has taught that there is a significant difference in differential expression for DMN as between a population of individuals having osteoarthritis and a population of individuals not having osteoarthritis, and further has taught to compare the level of expression of DMN in a test individual with populations having osteoarthritis and populations not having osteoarthritis using classification methods to determine the similarity or difference in gene expression levels as between the test subject and the tested populations (see paragraphs [0127] to [0128] and [0129] to [0137] in addition to paragraphs [0390] to [0393]. Independent claims 49, 68 and 71 require that the level of expression of RNA corresponding to DMN be compared with the level of DMN in other individuals who have osteoarthritis and require at a minimum a statistically significant similarity as between the test subject and control subjects having osteoarthritis before the level of gene expression of DMN is considered to be indicative of osteoarthritis, to be indicative of the test subject being a candidate for osteoarthritis, or to be classified with that of control subjects classified as having osteoarthritis, respectively.

Furthermore, the Applicant contends that it does not require undue experimentation for one of skill to determine the inherent direction or level of the statistically significant differential expression required for the claimed methods of detecting a osteoarthritis, given the widely established and validated analytical tools for analyzing gene expression levels. Therefore, it is not necessary for the Applicant to have taught the exact direction or level of difference between the two populations for one of skill to practice the invention. The Applicant has provided

sufficient information by teaching that DMN is differentially expressed and that the differential expression between healthy and control subjects is significant as between the populations.

The rejection particularly contends, at page 8, that if one were to detect expression of DMN in blood that is different from healthy patients, it would be unpredictable if this difference is due to the presence of osteoarthritis or to some other disease or condition, such as renal cell carcinoma, on the grounds that Twine *et al.* teaches that DMN is a gene that is differently expressed in renal cell carcinoma patients likely to progress to renal disease, as indicated at "*Table 4, ¶0307 and Example 4*".

Applicant respectfully traverses this contention on the grounds that the data set forth by Twine *et al.* relating to differential gene expression between such subjects (Tables 2A and 2B) is in fact clearly drawn to subject matter which is not encompassed by the claims. Namely, the data set forth by Twine *et al.* only relates to genes which are differently expressed in peripheral blood mononuclear cells (PBMCs or "mononuclear cells" ; i.e. lymphocytes and monocytes), hence in leukocytes which are fractionated into cell types. This is clearly evidenced, for example, in the materials and methods section of Twine *et al.* at Example 1 describing RNA sample isolation, in accordance with the recitation: "*PBMCs from the clinical trials were isolated from whole blood samples (8mL) collected into CPT tubes according to the standard procedure... PBMCs were purified over Ficoll gradients, washed two times with PBS and counted. Total RNA was isolated from PBMC pellets...*", and at Example 4, in accordance with the recitation "*[0598] Expression profiling analysis of the 20 disease-free PBMC RNA samples and 45 RCC PBMC RNA samples revealed that of the 12,626 transcripts on the HgU95A chip, 5,249 transcripts met the initial criteria for further analysis... On average, 4023 transcripts were detected as "present" in the 45 RCC PBMCs, while 4254 expressed transcripts were detected as "present" in the 20 disease-free PBMCs.*". In important contrast, the claims are drawn to RNA of samples which comprise unfractionated leukocytes, i.e. which comprise granulocytes as well as the minority mononuclear cell fraction of leukocytes. Thus, in setting forth data relating only to genes which are differentially expressed in a minority fraction of leukocytes, Twine *et al.* fails to provide teachings relating to genes which are globally differentially expressed in unfractionated leukocytes, as required by the claims. Namely, unfractionated leukocytes further comprise a

majority fraction of granulocytes which are a distinct cell type relative to mononuclear cells and which inherently have distinct gene expression profiles relative to mononuclear cells [refer, for example to enclosed abstract of Hashimoto S. et al., 2003. Gene expression profile in human leukocytes. Blood 101:3509-13; and to Figure 5C of the instant specification which indicates significantly different expression levels of an exemplary gene between granulocytes ("G.R.") and mononuclear cells, the latter being represented by cumulative levels of the combination of B-lymphocytes ("CD19"), T-lymphocytes ("CD3") and monocytes ("MONO")]. Thus, it cannot be predictably extrapolated that the global differential expression of any given gene (such as DMN between healthy control subjects and renal cell carcinoma patients) observed in Ficoll-fractionated mononuclear cells, as taught by Twine *et al.*, will also be observed in unfractionated leukocytes, as required by the claims.

As such the teachings of Twine *et al.* do not teach that the field of the invention is highly unpredictable, and cannot be used as grounds for contending that the claims are not enabled.

The rejection states that DMN is not differently expressed in subjects having osteoarthritis as well as asthma, obesity, hypertension or allergies, or having been subjected to systemic steroids. Applicant respectfully traverses this ground of rejection. Nevertheless, in the interest of expediting prosecution of the instant application, Applicant currently elects to file new independent claims 59, 68 and 71, which are drawn to a method of detecting DMN expression, a method of screening a human test subject for being a candidate for having osteoarthritis, and a method of classifying expression of a desmuslin (DMN) gene in a human test subject. Applicant additionally currently elects to add new claims 76 and 77 limiting the control subjects to those which are not subject to systemic steroids, or which do not have rheumatoid arthritis, hypertension, obesity, allergies, mild osteoarthritis, and/or severe osteoarthritis

As stated in the Manual of Patent Examining Procedure at 2164.03: the "predictability or lack thereof" in the art refers to the ability of one skilled in the art to extrapolate the disclosed or known results to the claimed invention. In the instant application, the disclosed result is a statistically significant differential expression in the level of DMN RNA as between subjects having osteoarthritis and subjects not having osteoarthritis, the statistically significant difference having a p value less than 0.05, as indicated in Table 3O of the instant specification.

The Office Action states on page 7 that Lee teaches that data obtained from gene chips must be replicated in order to screen out false positive results; that Cheung et al. (2003) teaches that there is natural variation in gene expression amongst different individuals (page 10 of the office action); that Wu et al (2001) teaches that gene expression data, such as microarray data, must be interpreted in the context of other biological knowledge, and that the conclusions that can be drawn from a given set of data depend on the particular choice of data analysis, (page 10 of the office action); and that Newton et al. (2001) teaches that a replication of data is required for validation (also page 10 of the office action).

Applicant respectfully disagrees with the contention based on Wu et al. that expression data needs to be interpreted in view of other biological knowledge. Differential gene expression which is reproducible, and is correlated with the state of health or disease of the individual does not necessarily result directly from the state of disease of the individual. Rather these changes in expression may simply represent a downstream side-effect of pathogenic processes, and it is not necessary that the biological relevance of the data be known to allow this difference in expression to be useful as a biomarker. For example prostate-specific phosphatase and prostate-specific antigen (PSA) were long used as biomarkers without an understanding of their function (refer, for example, to the enclosed abstracts of: Chu TM, 1990, Prostate cancer-associated markers. Immunol. Ser. 53:339-56; and Diamandis EP., 2000, Prostate-specific antigen: a cancer fighter and a valuable messenger? Clin Chem. 46:896-900).

The Examiner also argues, on the basis of post-filing art of Wu (2001) and Newton (2001), that many factors may influence the outcome of the data analysis and notes that conclusions depend on the methods of data analysis. While considerations such as variability, and normalization are of importance, these considerations are well understood by a person skilled in the art and have been applied for many years to permit development of biomarkers which are indicative of disease. These challenges are well understood, as are the routine experiments required to exemplify statistically significant differences in populations.

Applicant notes that the results disclosed by Cheung *et al.* cannot be reliably extrapolated to primary blood samples since the lymphoblastoid cells employed by Cheung *et al.* are significantly modified relative to primary blood cells, due to being cultured cell lines generated

by immortalization of primary human cells derived from "CEPH" families, as indicated in Reference no. 10 of Cheung *et al.* (Dausset *et al.*, 1990. Genomics 6:575; enclosed) at page 575, right column, 1st paragraph. Applicant notes that immortalized cultured cell lines such as the lymphoblastoid cells taught by Cheung *et al.* undergo significant genetic modification such as strong genome-wide demethylation (refer, for example, to enclosed abstract of: Vilain *et al.*, 2003. DNA methylation and chromosome instability in lymphoblastoid cell lines. Cytogenet Cell Genet. 90:93), as a result of extensive *in-vitro* culturing in the absence of immune or apoptotic mechanisms which function to eliminate mutated cells in the body. As such, immortalized CEPH lymphoblastoid cells may represent a particularly unsuitable cell type for modeling gene expression variability in primary blood cells.

To the extent that Cheung et al. could still be considered to suggest that larger populations of diseased and control populations may be useful to determine what level of differential expression is indicative of disease amongst the population at large, the Applicant submits that the extension of the experiments as outlined in the specification to additional individuals is merely routine. As is noted in Re Wands "*even a considerable amount of experimentation is permissible to practice the claimed methods, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed.*" (Re Wands, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988)).

Furthermore, the decision *In re Angstadt*, 190 U.S.P.Q. 218 (C.C.P.A. 1976) clearly states that even in an unpredictable art, and clearly permits the presence of a screening step to identify those embodiments which possess the desired activity is permissible. In fact, in *Angstadt*, the Court specifically dismissed the notion that the specification must provide a level of guidance that would predict the outcome of an experiment "with reasonable certainty before performing the reaction" and that "such a proposition is contrary to the basic policy of the Patent Act, which is to encourage disclosure of inventions and thereby to promote progress in the useful arts." The "predictability or lack thereof" in the art refers to the ability of one skilled in the art to extrapolate the disclosed or known results to the claimed invention.

Applicant wishes to point out that in *In re Wands*, the court stated that "[e]nablement is not precluded by the necessity for some experimentation such as routine screening. However, experimentation needed to practice the invention must not be undue experimentation. 'The key word is 'undue' not 'experimentation' (citing *In re Angstadt*, 537 F. 2d 498 at 504, 190 U.S.P.Q. 214 at 219 (C.C.P.A. 1976)). The Court also stated that "the test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed." (citing In re Jackson, 217 U.S.P.Q. 804 at 807 (Bd. App. 1982)).

As such the Applicants believe there is sufficient guidance provided by the specification that DMN gene is differentially expressed between human individuals who are healthy as compared to those having osteoarthritis, and that the art is sufficiently predictable such that the amount of experimentation to perform the instantly claimed methods of diagnosing osteoarthritis and identifying candidate subjects who may have osteoarthritis is not undue. In light of the amendments and above remarks, the Applicant contends that the claims are fully enabled, and respectfully requests reconsideration and withdrawal of the instant rejections.

Conclusion

Applicant submits that all claims are allowable as written and respectfully request early favorable action by the Examiner. No new matter is added. If the Examiner believes that a telephone conversation with Applicant's attorney/agent would expedite prosecution of this application, the Examiner is cordially invited to call the undersigned attorney/agent of record.

Respectfully submitted,

Date: November 30, 2007

Amy DeClaire 54849 for

Name: Kathleen M. Williams
Registration No.: 34,380
Customer No.: 21874
Edwards Angell Palmer & Dodge LLP
P.O. Box 55874
Boston, MA 02205
Tel: 617-239-0100

Encl:

Abstract of Casey et al., 1988. simplified plastic embedding and immunohistologic technique for immunophenotypic analysis of human hematopoietic and lymphoid tissues. Am J Pathol. 131:183-9;

Abstract of: Chu TM, 1990, Prostate cancer-associated markers. Immunol. Ser. 53:339-56;

Abstract of: Diamandis EP., 2000, Prostate-specific antigen: a cancer fighter and a valuable messenger? Clin Chem. 46:896-900);

Abstract of: Hashimoto S. et al., 2003. Gene expression profile in human leukocytes. Blood 101:3509-13;

Abstract of: Vilain et al., 2000. DNA methylation and chromosome instability in lymphoblastoid cell lines. Cytogenet Cell Genet. 90:93;

Dausset et al., 1990. Genomics 6:575;

Appendix A

1: Am J Pathol. 1988 May;131(2):183-9.

A simplified plastic embedding and immunohistologic technique for immunophenotypic analysis of human hematopoietic and lymphoid tissues.

Casey TT, **Cousar JB**, **Collins RD**.

Department of Pathology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232.

Routine fixation and paraffin embedding destroys many hematopoietic and lymphoid differentiation antigens detected by flow cytometry or frozen section immunohistochemistry. On the other hand, morphologic evaluation is difficult in flow cytometric or frozen section studies. A simplified three-step plastic embedding system using acetone-fixed tissues embedded in glycol-methacrylate (GMA) resin has been found to provide both excellent morphologic and antigenic preservation. With our system, a wide variety of antigens are detected in plastic sections without trypsinization or prolonged embedding procedures; pan-B (CD19, CD22), pan-T (CD7, CD5, CD3, CD2), T-subset (CD4, CD8, CD1, CD25) markers as well as surface immunoglobulin and markers for myeloid and mononuclear-phagocyte cells are preserved. In summary, modifications of plastic embedding techniques used in this study simplify the procedure, apparently achieve excellent antigenic preservation, and facilitate evaluation of morphologic details in relation to immunocytochemical markers.

PMID: 3282442 [PubMed - indexed for MEDLINE]

Appendix B

1: Immunol Ser. 1990;53:339-56.

Prostate cancer-associated markers.

Chu TM.

Roswell Park Memorial Institute, New York State Department of Health, Buffalo.

Immunodiagnosis of prostate cancer is at a more advanced stage than that of most other tumors. Two well-known markers, prostatic acid phosphatase and prostate-specific antigen, have been used in the clinical management of patients. Prostate-specific antigen is a more sensitive and reliable marker than prostatic acid phosphatase. Serum prostate-specific antigen is effective in monitoring disease status, predicting recurrence, and detecting residual disease. Prostate-specific antigen is a tool for the histological differential diagnosis of metastatic carcinomas, especially in the identification of metastatic prostate tumor cells in distant organs and in the differentiation of primary prostate carcinoma from poorly differentiated transitional cell carcinoma of the bladder. Few data on biological function are available. Prostatic acid phosphatase functions as a phosphotyrosyl-protein phosphatase and prostate-specific antigen as a protease. Physiological function in the prostate remains to be elucidated. Several of the prostate-specific and prostate-tumor-associated antigens, as well as a putative prostate tumor-specific antigen, as recognized by monoclonal antibodies are available. Clinical evaluation of these potential markers is not yet available.

PMID: 1713065 [PubMed - indexed for MEDLINE]

Appendix C

1: Blood. 2003 May 1;101(9):3509-13. Epub 2003 Jan 9.

Gene expression profile in human leukocytes.

Hashimoto S, Nagai S, Sese J, Suzuki T, Obata A, Sato T, Toyoda N, Dong HY, Kurachi M, Nagahata T, Shizuno K, Morishita S, Matsushima K.

Department of Molecular Preventive Medicine, Graduate School of Frontier Science, University of Tokyo, Japan.

Leukocytes are classified as myelocytic or lymphocytic, and each class of leukocytes consists of several types of cells that have different phenotypes and different roles. To define the gene expression in these cells, we have performed serial analysis of gene expression (SAGE) using human leukocytes and have provided the gene database for these cells not only at the resting stage but also at the activated stage. A total of 709,990 tags from 17 libraries were analyzed for the manifestation of gene expression profiles in various types of human leukocytes. Types of leukocytes analyzed were as follows: peripheral blood monocytes, colony-stimulating factor-induced macrophages, monocyte-derived immature dendritic cells, mature/activated dendritic cells, granulocytes, natural killer (NK) cells, resting B cells, activated B cells, naive T cells, CCR4(-) memory T cells (resting T(H)1 cells), CCR4(+) memory T cells (resting T(H)2 cells), activated T(H)1 cells, and activated T(H)2 cells. Among 38,961 distinct tags that appeared more than once in the combined total libraries, 27,323 tags were found to represent unique genes in certain type(s) of leukocytes. Using probability (P) and hierarchical clustering analysis, we identified the genes selectively expressed in each type of leukocytes. Identification of the genes specifically expressed in different types of leukocytes provides not only a novel molecular signature to define different subsets of resting and activated cells but also contributes to further understanding of the biologic function of leukocytes in the host defense system.

PMID: 12522010 [PubMed - indexed for MEDLINE]

Appendix D

☐ **1:** Clin Chem. 2000 Jul;46(7):896-900.

Prostate-specific antigen: a cancer fighter and a valuable messenger?

Diamandis EP.

Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada. ediamandis@mtsinai.on.ca

BACKGROUND: Prostate-specific antigen (PSA) is a valuable prostatic cancer biomarker that is now widely used for population screening, diagnosis, and monitoring of patients with prostate cancer. Despite the voluminous literature on this biomarker, relatively few reports have addressed the issue of its physiological function and its connection to the pathogenesis and progression of prostate and other cancers. APPROACH: I here review literature dealing with PSA physiology and pathobiology and discuss reports that either suggest that PSA is a beneficial molecule with tumor suppressor activity or that PSA has deleterious effects in prostate, breast, and possibly other cancers. CONTENT: The present scientific literature on PSA physiology and pathobiology is confusing. A group of reports have suggested that PSA may act as a tumor suppressor, a negative regulator of cell growth, and an apoptotic molecule, whereas others suggest that PSA may, through its chymotrypsin-like activity, promote tumor progression and metastasis. SUMMARY: The physiological function of PSA is still not well understood. Because PSA is just one member of the human kallikrein gene family, it is possible that its biological functions are related to the activity of other related kallikreins. Only when the physiological functions of PSA and other kallikreins are elucidated will we be able to explain the currently apparently conflicting experimental data.

PMID: 10894830 [PubMed - indexed for MEDLINE]

Appendix E

☐ **1:** Cytogenet Cell Genet. 2000;90(1-2):93-101.

DNA methylation and chromosome instability in lymphoblastoid cell lines.

Vilain A, **Bernardino J**, **Gerbault-Seureau M**, **Vogt N**, **Niveleau A**, **Lefrançois D**, **Malfoy B**, **Dutrillaux B**.

Institut Curie-CNRS UMR 147, Cytogénétique Moléculaire et Oncologie, Paris, France.

In order to gain more insight into the relationships between DNA methylation and genome stability, chromosomal and molecular evolutions of four Epstein-Barr virus-transformed human lymphoblastoid cell lines were followed in culture for more than 2 yr. The four cell lines underwent early, strong overall demethylation of the genome. The classical satellite-rich, heterochromatic,juxtacentromeric regions of chromosomes 1, 9, and 16 and the distal part of the long arm of the Y chromosome displayed specific behavior with time in culture. In two cell lines, they underwent a strong demethylation, involving successively chromosomes Y, 9, 16, and 1, whereas in the two other cell lines, they remained heavily methylated. For classical satellite 2-rich heterochromatic regions of chromosomes 1 and 16, a direct relationship could be established between their demethylation, their undercondensation at metaphase, and their involvement in non-clonal rearrangements. Unstable sites distributed along the whole chromosomes were found only when the heterochromatic regions of chromosomes 1 and 16 were unstable. The classical satellite 3-rich heterochromatic region of chromosomes 9 and Y, despite their strong demethylation, remained condensed and stable. Genome demethylation and chromosome instability could not be related to variations in mRNA amounts of the DNA methyltransferases DNMT1, DNMT3A, and DNMT3B and DNA demethylase. These data suggest that the influence of DNA demethylation on chromosome stability is modulated by a sequence-specific chromatin structure. Copyright 2000 S. Karger AG, Basel.

PMID: 11060456 [PubMed - indexed for MEDLINE]

Appendix F

PROGRAM DESCRIPTION

Centre d'Etude du Polymorphisme Humain (CEPH): Collaborative Genetic Mapping of the Human Genome

JEAN DAUSSET,* HOWARD CANN,* DANIEL COHEN,* MARK LATHROP,* JEAN-MARC LALOUEL,† AND RAY WHITE†,[1]

Centre d'Etude du Polymorphisme Humain (CEPH), 27 rue Juliette Dodu, 75010 Paris, France; and †Howard Hughes Medical Institute and Department of Human Genetics, University of Utah Health Sciences Center, Salt Lake City, Utah 84132

Elsewhere in this issue of *Genomics* is the first consortium map, that of chromosome 10, from the CEPH collaboration to map the human genome genetically. The map is truly a collaborative achievement, in that the underlying genotypes represent the efforts of laboratories collaborating with each other and with CEPH to produce a primary genetic map of the genome, consisting of polymorphic markers placed at approximately 20-cM intervals along each of the human autosomes and the X chromosome. Such a map provides a tool for the systematic localization of genes that determine inherited diseases and of other genes of interest. Genetic localization can be the first step in the development of diagnostic tests and isolation of a disease-determining gene. The purpose of this program description is to provide information about CEPH and the basis and nature of the collaboration.

CEPH

The Centre d'Etude du Polymorphisme Humain (CEPH)[2] is a nonprofit research institute that makes available to the scientific community a valuable research resource. CEPH is committed to (1) make available to the scientific community DNA samples from a panel of reference families for the determination of genotypes for various DNA polymorphisms which may be used for the construction of the genetic map of the human genome and for other research areas dependent on access to such a common set of families, and (2) provide to the contributors of genotypes a compilation of all data that accumulate on the panel of families.

CEPH Collaboration

The CEPH collaboration to map the human genome was organized in 1984 to hasten construction of a primary human genetic map with DNA polymorphisms (Botstein *et al.*, 1980). A key premise of the CEPH collaboration is that the human genetic map will be efficiently achieved by collaborative research on DNA from the *same* sample of families. To this end, CEPH provides to collaborating investigators high-quality cellular DNA produced from cultured lymphoblastoid cell lines (LCL) derived from each member of a reference panel of large nuclear families/pedigrees and a database contributed to and shared by these investigators. Collaborating investigators determine genotypes with their probes and the DNA from the CEPH panel to test the families for segregation of these genetic markers. They then contribute the genotypes to CEPH for preparation of a database which is returned to them for linkage analysis and map construction. As of October 1, 1989, 63 research laboratories in the United States (36), Canada (2), Europe (20), South Africa (2), Japan (2), and Australia (1) collaborate with CEPH in this manner.

CEPH Reference Family Panel

Families with large sibships, living parents, and grandparents are especially informative for linkage mapping (White *et al.*, 1985). From 100 families available from various sources, selected not for disease but for large sibship size, an initial group of 40 families was defined for the CEPH reference panel by the original group of collaborating investigators. Table 1 shows the geographic origins of these families and the contributors of the LCLs to CEPH. These are Caucasian families. The mean sibship size for these 40 families, based on those individuals for whom there are LCLs, is 8.3; no family has less than 6 offspring, and 23 families have 8 or more offspring. LCLs are available for all 4 grandparents in each of 29 families of the reference panel.

LCLs of the reference panel are stored in liquid nitrogen at three geographically separate repositories: Paris and Lyon, France; and Salt Lake City, Utah. These LCLs are not distributed by CEPH within or outside of the collaboration. They are used only as a source of DNA for the collaboration. LCLs for 11 of the panel families are available from the NIGMS Human Genetic Mutant Cell Repository located at the Coriell Institute for Medical Research in Camden, New Jersey.

Approximately 20 mg of DNA is prepared, as needed, from each LCL, and aliquots of 200 or 400 μg are distributed to each collaborating investigator. There is no charge to collaborating investigators for the DNA. The DNA is prepared by classical methods, scaled up to preparation of milligram amounts, which include lysis with proteinase K and SDS, extraction with phenol, and precipitation with isopropanol. The quality of each preparation of CEPH DNA is routinely controlled by testing for concentration, molecular size, digestibility with two different restriction enzymes, and contamination with vector sequences (Southern blot hybridized with a cosmid vector). In addition, each DNA preparation is hybridized with at least two highly variable ("minisatellite") probes (Wong *et al.*, 1987) to confirm identity of the LCL source and to detect contamination with DNA from another individual.

[1] All authors are members of the CEPH Executive Committee.
[2] Founded in 1983 by J.D. and D.C.

0888-7543/90 $3.00

TABLE 1

Sources of CEPH Family Panel

Utah	27 families[a]	R. White
France	10 families	J. Dausset
Venezuela	2 families[b]	J. Gusella
Pennsylvania (Old Order Amish)	1 family[c]	J. Egeland

[a] LCLs from 9 families available from NIGMS Human Genetic Mutant Cell Repository, Camden, NJ.

[b] LCLs from 1 of 2 families available from NIGMS Human Genetic Mutant Cell Repository.

[c] LCLs available from NIGMS Human Genetic Mutant Cell Repository.

DNA Polymorphisms

The family panel DNA is being tested for at least 1200 DNA polymorphisms. The probes being used within the collaboration detect polymorphism within restriction enzyme sites of low copy number or unique DNA segments, due to varying numbers of tandemly repeated, relatively small sequences, or VNTRs (Nakamura *et al.*, 1987), and of centromere-associated alpha satellite DNA (Willard *et al.*, 1986). Some collaborating investigators have begun to use genetic markers with family panel DNA that are based on detection of length variations of very small sequences amplified by PCR (Weber and May, 1989; Litt and Luty, 1989). Phenotypes for classical, serological, and electrophoretic polymorphic markers are also available for the family panel.

CEPH Database

CEPH collaborating investigators have agreed on two basic rules concerning their participation within the collaboration: (1) A CEPH collaborating investigator is committed to screening the reference panel by determining genotypes for all 40 parent pairs with each probe and enzyme combination used and then following segregation in each informative family (at least one heterozygous parent) in his or her own laboratory or collaboratively. (2) Collaborating investigators are committed to sending genotypic data to CEPH for inclusion in the CEPH database no later than publication of the data themselves or of an article based on the data.

Genotypic data generated from family panel DNA are processed into uniform format by a set of programs developed at CEPH (J.-M.L.) for use with IBM PCs or compatible models and sent to all collaborating investigators. The data files thus prepared are sent to CEPH to be merged into the CEPH database.

There are two components of the database: The CEPH collaborative database is available only within the collaboration. Two classifications of data are recognized in the collaborative database, *unpublished* and *published*. The first are those data that have not been used for a publication. These are privileged data, requiring permission from the contributor for use (e.g., publication) by another collaborating investigator. Published data are automatically available for inclusion in CEPH consortium maps (see below) and after a lag period will be released to the CEPH public database. After 2 years in the collaborative database, unpublished data become published data. The CEPH public database, currently being organized, will contain published data released from the collaborative database. Data in the public database will be available to the general scientific community.

Table 2 summarizes the contents of the CEPH collaborative database. As of July 1989, the database contains genotypic data for 1061 genetic markers localized to all the autosomes and the X chromosome (including the pseudoautosomal region). Approximately 20% of these marker systems have four or more alleles.

New CEPH Activities

As the primary map of the human genome nears completion, there is growing interest in a higher resolution genetic map, perhaps of the order of 1–2 cM. Availability of a higher resolution map will increase the efficiency and precision of localization of genes. In order to support construction of a high-resolution map of the genome, CEPH is in the process of increasing the family panel to 61 large nuclear families/pedigrees. LCLs from 21 additional families have already been received from Utah, and stocks of DNA are being prepared from them. The mean sibship size for the total of 61 families in the augmented panel is 8.5, and LCLs for all 4 grandparents are available for each of 44 families. These 21 families provide an additional advantage for genetic mapping in that they have already been genotyped for approximately 500 genetic markers used in the CEPH collaboration.

CEPH has undertaken a project to enhance the use of the primary map of the genome in localizing genes of interest. The idea is to collect and produce quantities of ready-to-label probes for the mapped primary markers and distribute these to biomedical scientists who wish to localize disease-determining genes and other genes of interest. Probes to be distributed will be those for markers chosen from CEPH consortium maps based on criteria that include heterozygote frequency and position on the genetic map of a chromosome. Collections of probes ("kits") will be available for each chromosome. By using the probes in a kit to test a large kindred or group of families in which a genetic disease with the appropriate chromosome assignment is segregating, an inves-

TABLE 2

CEPH Database

Version	No. of contributing labs	No. of markers	No. of markers with 4 or more alleles	Date
V1	17	171	28	April 1987
V2	29	744	143	March 1988
V3	33	1061	204	July 1989

tigator should be able to find linkage to one, two, and perhaps more markers if a sufficient number of informative meioses are available. Again, if data from a sufficient number of meioses are available, we would expect that a gene of interest without a chromosome assignment could be localized with markers from all of the kits. Investigators using these kits for primary mapped markers will be invited to contribute the genotypes they determine from families being tested with the probes to a database provided by CEPH. This project, being carried out in collaboration with the American Type Culture Collection, is sponsored by the National Institutes of Health.

Progress toward the Primary Genetic Map of the Human Genome

The number of markers for which there are genotypic data in the CEPH database suggests that the primary human genetic map is nearing completion. Partial or nearly complete[3] primary linkage maps based on genotypes determined from reference panel DNA with probes from a single or small group of CEPH collaborating laboratories have been published for many of the chromosomes (see, for example, Donis-Keller *et al.*, 1987; O'Connell *et al.*, 1989; Lathrop *et al.*, 1988; Nakamura *et al.*, 1988; Warren *et al.*, 1989). CEPH consortium maps, the first of which appears in this issue of *Genomics*, will provide reference primary linkage maps of each chromosome based on the combined genotypic data within the collaboration.

CEPH Consortium Maps

The primary genetic map of the human genome, based on all genotypes determined from the CEPH families, will be communicated to the scientific community through a series of consortium maps, one for each chromosome. These maps will represent the ultimate validation of genotypic data in the CEPH database. Each consortium map will be prepared by a committee of collaborating investigators who have contributed genotypic data for markers localized to the particular chromosome. A chromosome-specific database, containing genotypes for all relevant markers, will be sent to members of the consortium map committee for the construction of genetic maps. These maps will be circulated within the committee for study and comparison. The committee will meet for the final analysis of these maps, which will be used in the preparation of the consortium map. The consortium map and a report will be published in *Genomics*. It is this process that led to the CEPH consortium map of chromosome 10 presented in this issue of *Genomics*.

As each consortium map is published, the underlying genotypic data will be released to the scientific community. The consortium maps will provide the basis for choosing the primary mapped markers for which kits of probes will be distributed by CEPH to the scientific community in order to enhance the use of the genetic map for localization of genes that determine disease and other genes of interest.

[3] A primary linkage map of a chromosome cannot be considered complete until it contains markers for, or quite close to, the telomeres.

REFERENCES

1. Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map using restriction fragment length polymorphisms. *Amer. J. Hum. Genet.* **32**: 314–331.
2. Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T., Bowden, D., Smith, D., Lander, E., Botstein, D., Akots, G., Rediker, K., Gravius, T., Brown, V., Rising, M., Parker, C., Powers, J., Watt, D., Kauffmann, E., Bricker, A., Phipps, P., Muller-Kahle, H., Fulton, T., Ng, S., Schumm, J., Braman, J., Knowlton, R., Barker, D., Crooks, S., Lincoln, S., Daly, M., and Abrahamson, J. (1987). A genetic linkage map of the human genome. *Cell* **51**: 319–337.
3. Lathrop, M., Nakamura, Y., O'Connell, P., Leppert, M., Woodward, S., Lalouel, J.-M., and White, R. (1988). A mapped set of genetic markers for human chromosome 9. *Genomics* **3**: 361–366.
4. Litt, M., and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Amer. J. Hum. Genet.* **44**: 397–401.
5. Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**: 1616–1622.
6. Nakamura, Y., Lathrop, M., O'Connell, P., Leppert, M., Barker, D., Wright, E., Skolnick, M., Kondoleon, S., Litt, M., Lalouel, J.-M., and White, R. (1988). A mapped set of DNA markers for human chromosome 17. *Genomics* **2**: 302–309.
7. O'Connell, P., Lathrop, G. M., Nakamura, Y., Leppert, M. L., Ardinger, R. H., Murray, J. L., Lalouel, J.-M., and White, R. (1989). Twenty-eight loci form a continuous linkage map of markers for human chromosome 1. *Genomics* **4**: 12–20.
8. Warren, A. C., Slaugenhaupt, S. A., Lewis, J. G., Chakravarti, A., and Antonarakis, S. E. (1989). A genetic linkage map of 17 markers on human chromosome 21. *Genomics* **4**: 579–591.
9. Weber, J. L., and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Amer. J. Hum. Genet.* **44**: 388–396.
10. White, R., Leppert, M., Bishop, T., Barker, D., Berkowitz, J., Brown, C., Callahan, P., Holm, T., and Jerominski, L. (1985). Construction of linkage maps with DNA markers for human chromosomes. *Nature (London)* **313**: 101–105.
11. Willard, H. F., Waye, J. S., Skolnick, M. H., Schwartz, C. E., Powers, V. E., and England, S. B. (1986). Detection of restriction fragment length polymorphisms at the centromeres of human chromosomes by using chromosome-specific a satellite DNA probes: Implications for development of centromere-based genetic linkage maps. *Proc. Natl. Acad. Sci. USA* **83**: 5611–5615.
12. Wong, Z., Wilson, V., Patel, I., Povey, S., and Jeffreys, A. J. (1987). Characterization of a panel of highly variable minisatellites cloned from human DNA. *Ann. Hum. Genet.* **51**: 269–288.